

**DECLARATION OF ERIC BRENNER IN OPPOSITION TO PLAINTIFFS'
MOTION IN LIMINE TO EXCLUDE THE LITIGATION SURVEY BY
DEFENDANTS' EXPERT KEVIN LANE KELLER, PH.D.**

EXHIBIT 11

Demand Artifacts in Laboratory Experiments in Consumer Research

ALAN G. SAWYER*

Demand artifacts may influence the results of experiments in consumer behavior. After reviewing problems of demand artifacts, this article discusses four methods to investigate demand bias. Several suggestions involving experimental design, measures, and procedures to try to reduce demand problems in consumer research are presented.

SOURCES AND EFFECTS OF DEMAND ARTIFACTS

Demand artifacts or characteristics were first explicitly defined by Orne (1962). As he subsequently stated,

Insofar as the subject cares about the outcome, his perception of his role and of the hypothesis being tested will become a significant determinant of his behavior. The cues which govern his perception—which communicate what is expected of him and what the experimenter hopes to find—can therefore be crucial. . . . They include the scuttlebutt about the experiment, its setting, implicit and explicit instructions, the person of the experimenter, subtle cues provided by him, and, of particular importance, the experimental procedure itself. All of these cues are interpreted in the light of the subjects' past learning and experience. Although the explicit instructions are important, it appears that subtler cues from which the subject can draw covert or even unconscious inference may be still more powerful (Orne, 1969, p. 146).

Thus demand characteristics include all aspects of the experiment which cause the subject to perceive, interpret, and act upon what he believes is expected or desired of him by the experimenter. Artifacts such as suspiciousness of the experimenter's intent, the subject's original willingness to participate, the subject's past experimental experience, and obtrusive pre- and post-treatment measurement can produce demand bias if these artifacts increase the possibility that the subject believes he has encoded the true purpose of the experiment or if the artifacts affect his perceptions of appropriate behavior based upon his discovery (see Rosenthal and Rosnow, 1969; Rosnow and Aiken, 1973).

The effects of demand artifacts depend upon adopted

subject roles. Orne believed that most subjects are motivated to try to confirm what they believe to be the experimental hypothesis. The positiveness of science and the dependence of the researcher on the subject were believed to be sufficient motivators of this *good* role. However, there are at least three other possible subject roles (Weber and Cook, 1972). The *faithful* subject (Fillenbaum, 1966) is one who believes in near complete docility in research settings, is concerned only with following instructions scrupulously, and avoids acting on the basis of any suspicions he might have about the true purpose of the study. Bias should not result from this role. The *negative* subject role occurs when a subject tries to disconfirm a suspected experimental hypothesis by behaving in a contrary or intentionally random or neutral manner (e.g., Agyris, 1968; Brehm, 1966; Masling, 1966). The effects of a fourth possible subject role—the *apprehensive subject* (Rosenberg, 1969; Riecken, 1962) are most ambiguous and unpredictable. Subjects often worry about how their performance in an experiment will be used to evaluate their abilities or socioemotional adjustment. The unpredictability of the consequences of this apprehensive role arises from the fact that a subject may misinterpret either the experimenter's hypothesis or the behavior that will reflect most admirably upon the subject himself.

The effects of demand artifacts pose important threats to both internal and external validity. Contaminating demand characteristics become a major problem of internal inference when they interact with the effect of the independent variable under study. Such interactions may cause false positive findings when confirmed hypotheses attributed to some theoretical notion are actually due to demand cues. Alternatively, demand cues may suppress a hypothesized relationship and lead

* University of Massachusetts.

DEMAND ARTIFACTS IN CONSUMER RESEARCH

21

to false negative findings. The latter discouraging result can also be produced by diverse adopted subject roles which inflate error variance and decrease statistical power. Demand characteristics which are more likely to occur in the more artificial laboratory may affect the ability of the experimenter to generalize his results to a real life situation where an analogous set of demand conditions may be absent (e.g., Silverman, 1968).

DEMAND ARTIFACTS IN CONSUMER BEHAVIOR

Many consumer behavior experiments may be affected in some way by demand artifacts. Two research areas—quality connotations of price and repetition and attitude change—have seemed especially prone to demand bias.

Quality-Connotation of Price

A popular research topic in consumer behavior is the question of whether consumers utilize price information as a cue to quality or desirability. These studies often ask subjects to choose or rate products where the only available information is the product's price (e.g., Leavitt, 1954; Tull *et al.*, 1964; McConnell, 1968; Peterson, 1970; Deering and Jacoby, 1972). For example, Tull, Boring, and Gonsior (1964) studied the effects of price, product quality differences, and reference price. Subjects were presented with three different brands in each of four product categories. Two product categories were judged quite low in terms of interbrand quality differences and two were relatively high in this respect. Reference price was manipulated over three treatment groups by telling the subjects to assume that the price of the brand they usually bought was the same as either the low, medium, or high price associated with the brands. Two main results were reported: 1) where product differences were felt to be great, higher priced brands were more apt to be chosen than when differences were small, and 2) the most often chosen price was at the reference price. In a different procedure, McConnell (1968) presented subjects a choice between three brands of beer which were differentiated only by an identifying letter and the relative price (the actual beers were exactly the same). Over twenty-four weekly trials, subjects were most apt to choose the high-priced beer, but more apt to choose the low priced brand than the medium one.

Demand artifacts could account for the above results. The subject usually perceives that some logical means of performing his experimental task is available. For subjects making any attempt at all to decipher the purpose of the experiment, price as the only piece of information available would certainly be a prime aspect on which to focus. In the first study, subjects might rationally hypothesize that the experimenter expected greater choice of higher priced brands primarily for

products which differ in quality. Furthermore, by choosing the brand at the reference price given by the experimenter, the subjects were, in effect, acting exactly as they were told they usually did. With no explicit reference point in the second study, subjects focusing upon price as a cue to the experimenter's expectations would simply have to decide whether a positive or negative relationship was intended. Varying decisions between these two alternative behaviors would result in the observed curvilinear effect of price on brand choice.

In a survey of all published price-quality experiments, Olson (1973) found a significant main effect of price in six of six experiments where price was the only informational cue, but in only seven of twelve studies where information in addition to price was available. Such results may indicate that the other factors actually mediate the effects of price. However, a demand effects explanation would assert that the decreased effect of price is due to an increased difficulty of the subjects to perceive and act upon the likely experimental hypothesis.

Results of research that offers empirical evidence about such hypothesized demand cues will be reported in this paper. At this point, however, the tendency for a subject to seize upon some criterion for performing the experimental task should be emphasized. Peacock and Davis (1969) contended that McConnell's price-quality result could be alternatively explained by the confounding letter-brands identifying the low (P), medium (L), and high (M) priced beers. Their experiment which merely asked subjects to choose among those three letters and which did not mention beer or price almost identically replicated McConnell's results and appeared to leave his cited cause-and-effect relationship open to interpretation. Although a later experiment by McConnell (1970) with a factorial design varying both the L, M, and P alphabet "brands" and the price of the beer indicated that the price cue dominated the alphabet cue, Peacock and Davis' experiment dramatically demonstrated how subjects will seize upon any available cue to try to perform an experimental task and, perhaps, to also try to understand the experimental hypothesis.

Repetition and attitude change

Several experiments have examined the effects of repeated stimulus exposures on attitudes (e.g., Becknell *et al.*, 1963; Heeler, 1972; Miller *et al.*, 1971; Ray and Sawyer, 1971; Sawyer, 1973; Sawyer, 1974b). These experiments typically present several levels of exposures of several stimuli to a subject and then rate his responses to either those stimuli or objects linked to those stimuli.

For example, Zajonc (1968) reported that increasing the number of exposures of nonsense syllables re-

sulted in increased liking for those stimuli. After telling subjects they were participating in a study of language, Zajonc exposed slides of the stimuli for two seconds each and varied the number of exposures up to twenty-five. After viewing the slides, the subjects were asked to rate the goodness of the syllables by indicating whether they thought the words meant something good or something bad. The positive relationship between repetition and affect has since been replicated by Zajonc and his colleagues several times. Becknell, Wilson and Baird (1963) studied the effect of repeated exposures on brand choice. After viewing a set of nonsense syllables with frequency levels as high as ten, each respondent was asked to choose a gift of nylon stockings from two brands, each of which was marked by one of the repeated nonsense syllables seen in the presentation. A second showing was followed by a paper-and-pencil ranking of four "brands." Both measures were positively affected by repetition. Miller, Mazis, and Wright (1971) also found positive effects of as many as twenty exposures of different nonsense syllable words. After viewing slides of exposures of four words at the different exposure levels, subjects were told that the words were brand names for a new product. Each of the brand names was then linked to various types of form, color, and price characteristics. When subjects rated their liking for those characteristics, liking was positively related to the number of exposures of the nonsense syllable brand name linked to those characteristics. However, no effect of repetition on liking of taste was found when subjects were also able to actually sample the product with which the repeated brand name was linked.

Other research has failed to replicate the positive repetition-affect relationship (Ray and Sawyer, 1971; Heeler, 1972; Sawyer, 1973). These experiments used an experimental cover both to disguise the true purpose of the experiment and to try to make the repetition of advertisements seem plausible. Ray and Sawyer (1971) repeated slides of actual print ads up to six times for three brands within six product classes. Although brand purchase intention was positively affected by repetition, there was no significant effect of repetition on brand evaluation. Moreover, when the rating of the brand was compared to the rating of competing brands in order to develop a score of preference to other brands in the product class, the effects of repetition coincided with a significant decrease in preference.

It seems that positive effects of repetition on affect are more likely to result from experimental procedures in which the repeated exposures are presented with little or no accompanying "cover" or rationale for the subject and where no information other than the experimentally manipulated number of stimulus exposures is available to the subject. Inconsistent past results might be explained by differences in the environment of the experiments and in the presence of demand char-

THE JOURNAL OF CONSUMER RESEARCH

acteristics which may cue a subject to respond to repeated stimuli. An examination from a subject's viewpoint of the typical experimental procedure utilizing nonsense syllables as stimuli reveals that his role in the experiment might be perceived as quite strange and artificial. The task of rating or choosing among the nonsense syllables is a near impossible one. However, a subject realizes that the experimenter is performing the experiment from some purpose and expects the subject to rate the stimuli in some meaningful way. Thus, many subjects are likely to search for any available cues to differentiate the stimuli. The major difference in the stimuli is that, for some unexplained reason, the words were repeated several times but at different exposure frequencies. The fact that no warning or rationale for that repetition was offered by the experimenter adds to the artificiality of the experimental setting and may further focus the subject's attention to that variable. A subject who does not reject the task as too difficult or inane is likely to seize upon the widely manipulated number of exposures to differentiate his ratings on the stimuli. To the extent that the subject perceives that the experimenter expects some use of the exposure variable in the rating task and to the extent he believes the relationship should be positive, the subject may increase his ratings for stimuli that were exposed more often.

Stang (in press) has found that repetition experiments with an interval between the repeated exposures and affect rating are much more likely to produce significant effects of repetition than experiments with no intervening period. He speculated that "hypothesis guessing" occurs during these intervals whereas such guessing is minimized in the absence of an interval and the "true relationship" between repetition and affect emerges. Research discussed later in this paper indicates that subjects who take part in procedures like those of Zajonc do try to guess the experimental hypothesis in repetition experiments and that they tend to behave as they believe they are expected. Although negative and curvilinear hypotheses have also been expressed for most types of stimuli (Stang, 1974) a majority of subjects who do include repetition in their estimates of the experimenter's hypothesis equate repetition with goodness.

It also has been found that positive connotations by the subject of the experiment, the experimenter, and the subject's experience are correlated with a positive relationship between repetition and affect (Burgess and Sales, 1971). The fact that experiments with advertisements as repeated stimuli more often yield neutral or negative effects of repetition might be due to subjects who negatively evaluate the social and aesthetic values of advertising (Bauer and Greyser, 1968) and who are more apt to adopt a "negative" or "apprehensive" subject role or to feel negative about their participation in the experiment. Currently, however, there is no evi-

DEMAND ARTIFACTS IN CONSUMER RESEARCH

23

dence to support such hypothesized subject negativism. Demand artifacts might explain the fact that brand evaluation or attitudes are often not affected by repeated ad exposures but measures of purchase intention are positively affected (see Sawyer, 1974b). Subtle demand bias might affect the simpler purchase intention response of simply checking a brand but might be not strong enough to significantly alter a subject's relative ratings of several brands in a product class.

EVALUATION OF POTENTIAL DEMAND ARTIFACTS

There are four research modes that can examine the validity of such conjecture about demand artifacts: 1) the post-experimental inquiry, 2) the non-experiment, 3) manipulation of suspected demand cues, and 4) hetero-method replication. Investigations of these types have focused on the above two areas of consumer research and others as well. This research is presented both to expand on the possible demand bias in these areas and to illustrate the use of these methods.

Post-experimental inquiry.

Upon completion of a subject's participation in an experiment, it should be standard procedure to interview the subject (Aronson and Carlsmith, 1968; Tybout and Zaltman, 1974). Such a practice takes advantage of the fact that human subjects can talk and reflect upon their experience. The inquiry may include a debriefing which explains the purpose of the research, the subject's contribution, and the reason for the procedure including deception if used.¹ A prime purpose of the post-experimental interview should be to determine any suspicions the subject may have developed about the purpose of the experiment and the subject role he might have adopted. If inquiries reveal demand bias and if contaminated subjects can be satisfactorily isolated, these subjects can be dropped or statistically accounted for in the final analysis. Appropriate inquiries yielding no signs of demand bias can be presented to help refute speculation to the contrary.

Post-experimental inquiries in experiments using the Zajonc paradigm have offered support for a demand characteristics explanation. Sawyer (1974a) concluded a repetition experiment by asking subjects open-ended questions about their perceptions of the experiment, reasons for their behavior, and the experimenter's hypothesis. A minority of subjects who consciously used repetition as a rating criterion accounted for the majority of the observed positive repetition-affect relationship. Answers about subjects' perceptions of the ex-

¹ Of course, post-inquiries may help to reveal artifacts other than demand characteristics. Peterson (1970) sampled about fifteen percent of his subjects in a price-quality experiment and found that quality ratings depended upon the quality of reference product with which the subject assumed he should compare the experimental new product.

perimenter's hypothesis revealed that the fifteen percent who perceived a repetition-equals-good hypothesis responded extremely positively to repeated words while the ten percent who perceived a hypothesis of negative connotations of repetition reacted very negatively to repetition. The results lent strong support to the contention that subjects tend to adopt a "good" role and try to confirm what they believe to be the experimenter's hypothesis. Burgess and Sales (1971) found a similar pattern of response depending upon whether subjects reported positive or negative feelings about their participation in the experimental context.

Postexperimental interviewers have learned that accurate and honest feedback is far from assured. Of particular concern is the fact that subjects often are unwilling to disclose the true extent of their suspicions (Levy, 1967; Golding and Lichenstein, 1970). Often, in lab experiments, there exists an implicit "pact of ignorance" in which the subject does not wish to disclose anything because he will then have wasted his and the experimenter's time and/or because he correctly or incorrectly perceives that the interviewer does not really want his study delayed or disqualified. Such problems are especially salient in experiments which attempt to deceive the subject in some way, where declared suspicions involve calling the experimenter a liar.

In addition to unwillingness, inability to disclose perceived demand cues can pose a problem. Just as a subject may be unaware of independent treatments and yet be affected by them,² demand cues can have an unconscious effect. Moreover, subjects may have great difficulty articulating barely discerned demand aspects. Such a problem may be especially acute when it is noted that often subtle cues are more potent than more explicit ones. On the other hand, reports of demand cues may be overstated to the degree that the subject's perception of such cues evolved after the experiment.

Disclosure of subject perceptions and suspicions may be encouraged by use of an interviewer other than the experimenter. Whoever performs the task must assure the subject that he truly desires information about any suspicions from the subject. Questions could probe estimates of the experimental hypothesis, suspicions about any deception, perceived definitions of both good and apprehensive subject role behavior, and what conscious behavior criteria were employed. However, the interviewer must take care not to over-cue a subject since there may be a tendency for a subject to hesitate admitting he was the victim of a hoax or that he was unaware of any experimental purpose.

² LoSciuto, Strassman, and Wells (1967) found in a repetition-choice experiment that, even though most subjects appeared to be highly influenced by the repetition level, subjects claimed to have been unaware of the repetition and uninfluenced by it.

For example, Stang (1974) asked subjects to check the statement that best expressed the subjects' estimate of the experimenter's hypothesis. Five explicit hypotheses (including four involving repetition) plus a miscellaneous category were stated. Although such explicit wording of hypotheses may have helped subjects to articulate their opinions, such a practice may have artificially prompted subjects to think about repetition for the first time.

As an aid to determining if an experimenter can ascertain whether a subject is honestly responding to post-experimental questions, subject simulators (Orne, 1969) can be employed. These simulators pretend to have been affected by an experimental treatment in the presence of an experimenter blind to the subject's status as a pretended subject. Despite the measurement problems involved in post-experimental inquiries, some attempt to measure subjects' reactions to an experiment should be included in every laboratory investigation.

Non-experiment

A non-experiment involves a reenactment of all experimental procedures except the actual treatment, which is instead only described. New subjects from the same population sampled in the experiment in question are asked to role-play the subjects in the actual experiment. They undergo the same experimental procedure as the actual subjects. They are shown the same room and the employed equipment, read the same instructions, and, after the experimental treatment is described, asked to produce data just as if they had actually undergone the experimental treatment. A comparison of results from the non-experiment and the actual experiment can offer insights into the role of hypothesized demand artifacts in the procedure, design, or measurement instructions. Although similar results do not prove the presence of bias due to demand characteristics and may instead merely indicate that subjects are good predictors of their behavior, such results do demonstrate that a demand characteristic explanation is a plausible rival hypothesis, and further investigation is warranted. In addition, the cooperative, role playing subject may provide additional insight into the role of potential demand cues by reporting why he would behave as he describes.

Sawyer (1974a) used the non-experiment to investigate the possibility of demand cues in the Zajonc (1968) repetition and attitude change experimental procedure. Subjects were confronted with the same experimental instructions and measurement tasks, except that the repetition treatment was not actually experienced. The multiple exposure of various symbols was described by indicating on the questionnaire the number of times each symbol would have been exposed had the subjects actually viewed the stimuli. No cues, other than those present to any individual in the pro-

cedure of the original experiment, were given about the experimental hypothesis. The result of the non-experiment almost exactly matched Zajonc's positive repetition-affect result where repeated exposure was actually experienced.⁸ The fact that similar results were obtained from merely experiencing the experimental procedure demonstrated that the experimenter's hypothesis was not counter-intuitive to the subject's. In addition, the non-experiment results indicated that a demand characteristics explanation of Zajonc's results remained a plausible rival hypothesis—especially when the previously discussed answers to the post-experimental inquiry were also considered.

Brehm and Cohen (1959) asked young children to choose a toy and then studied the effects of the choice on subsequent ratings of both the chosen product and rejected alternatives. The results were that 1) the ratings of the chosen product increased while ratings of the rejected product decreased, and 2) this displacement effect was directly proportional to both the number of alternatives and the dissimilarity of the competing alternatives. These results, which have been explained by differences in the amount of aroused cognitive dissonance, have been replicated several times in other consumer behavior experiments (e.g., LoSciuto and Perloff, 1967; Anderson *et al.*, 1966; Mittelstaedt, 1969). Bem (1967) used the non-experiment procedure to test the plausibility of his alternative hypothesis that the displaced ratings were due to the subjects' simple deductions about their attitudes derived from observing their own choice behavior. The results of the non-experiment, which matched those of the original experiment, demonstrated that cognitive dissonance was not necessary to explain the results and that the experimental post purchase behavior was not counter expectational to the subjects. The successful replication in the non-experiment was also consistent with a demand cues hypothesis that the subject who chose a product in view of the experimenter would perceive that the experimenter, as an outside observer, expected him to change his ratings based on that choice.

Heteromethod replication

Multiple operationalism of independent variables and accompanying procedures can answer questions about confounding demand cues. If more than one procedure yielded similar results, the result (and accompanying theoretical explanation) would obviously carry more credibility.

For example, Craig *et al.*, (1971) compared two methods of investigating the effects of price changes on brand choice. One method was a verbal interview in which subjects were more likely to be highly conscious of prices and cued to switch brands when prices

⁸ Stang (1974) reported results nearly identical to Sawyer's non-experiment and post-experimental inquiry.

DEMAND ARTIFACTS IN CONSUMER RESEARCH

25

were changed; brand choice was measured on paper-and-pencil scales both before and after prices were changed until a subject altered his brand choice. The second method was a less demand-prone one in which great care was made not to sensitize the subject to price; prices were changed only once and the post-price change measurement was obtained from a simulated shopping trip unconnected with the earlier pre-rating. Comparison of the results in the two procedures showed that, of the 50% correctly predicted brand choices based on changed prices in the former experimental setting, less than half could be predicted in the potentially less demand biased procedure. More such heteromethod replication studies are needed in consumer research.

Manipulation of demand characteristics

When demand characteristics are suspected, experimental variation of the hypothesized cues can help to explore their effects. For example, subject's suspiciousness of the experimenter's intent can be experimentally varied (McGuire, 1969; Horowitz and Rothschild, 1970) or subject role set can be altered (Sigall *et al.*, 1970; Adair and Schachter, 1972; Rosnow, *et al.*, 1973). Ideally, a type of Solomon four-group design (Campbell and Stanley, 1963) in which the suspected demand cue is or is not included in the experimental treatment would be employed.

Regarding the effects of repeated exposure on attitudes, the Zajonc research paradigm has prompted some experiments which manipulated suspected demand cues. One such experiment examined the hypothesis that Zajonc's subjects reacted positively to repetition due to measurement instructions cueing a positive reaction. Suedfeld *et al.* (1971) manipulated the direction of measurement "set" as either positive or negative ("tell me the extent that this symbol means something good [bad]"). They found that a good set coincided with positive monotonic effects of repetition on attitude whereas a negative set led to an inverted U-curve relationship.

Burgess and Sales (1971) hypothesized that classical conditioning between positive evaluation of participation in the experiment and high frequency stimuli might account for Zajonc's positive results. In an experiment that manipulated both frequency of nonsense words and the positiveness of paired-associated words (very negative, negative, positive, very positive), the relationship of rated liking of the nonsense words at different frequency levels varied with the context of the associated words. A negative relationship was found for the negative context conditions, a positive relationship was found for the positive conditions, and the rank order of the positiveness of response to repetition exactly matched the hypothesized order of the four context conditions. Thus it would appear that subtle demand

cues in the measurement instructions could account for the positive repetition-attitude relationship. Some negative support for an experimental context demand artifact explanation was reported, however, by Saegert, Swap, and Zajonc (1973) in a study of repeated exposures on liking for another person. Even in instances of very unpleasant experimental surroundings in which tastes and smells were quite odious, a positive repetition-liking relationship was observed.

As part of an experiment related to the price-quality question, Deering and Jacoby (1972) assessed the effects of possible demand characteristics in the wording of one measure of price limits. They hypothesized that prompting subjects to indicate limits with the category labels of TOO EXPENSIVE TO BUY and TOO CHEAP TO BUY would result in more narrow limits than the labels, "very low in cost" and "very high in cost." However, no differences in the two conditions were found.

Some cautions about experiments deliberately manipulating demand characteristics are appropriate. First, the logic of manipulating suspicion of intent or deception implies monotonic effects. However, as noted, subtle cues may have more influence than more blatant ones. Second, in experiments manipulating subject roles, it is often difficult to satisfactorily provide an unambiguous behavior outlet for the multi-faceted apprehensive subject role (Weber and Cook, 1972).

REDUCING DEMAND BIAS

A researcher can never be absolutely sure about the presence of demand characteristics in his chosen experimental procedures. As Agyris stated, such contamination is inevitable. "The issue therefore is not contamination versus no contamination. The issue is under what conditions can the researcher have the greatest awareness of, and control over, the degree of contamination" (Agyris, 1968, p. 194). This last section offers several suggestions to reduce demand characteristics and their potential confounding effects in consumer behavior experiments. These suggestions relate to research design, measurement, and procedure.

Experimental designs

Two issues involving the source of experimental variation have potential relevance to demand bias. The first issue involves the use of between-subject versus within-subject designs. Within-subject designs involve the administration of a number of experimental treatments to the same subject who acts as his own control. Although such a design is efficient statistically and cost-wise and may be the only one possible in some cases (e.g. learning), it seems more demand-prone than a between-subjects design in which an individual subject is exposed to only one treatment level. The more conditions to which a subject is exposed, the more likely

he is to successfully estimate the experimenter's hypothesis. Perhaps reflecting this to some degree, experimental effects in within-subject designs are commonly of greater magnitude than in between-subject designs (Grice, 1966).

Regarding the within-subject problem in price-quality experiments, Olson (1973) found that eleven of thirteen experiments employing within-subject price manipulations found significant main effects of price whereas no main effect of price was found in three between-subject designs. For example, Andrews and Valenzi (1971) varied three levels each of price, brand name familiarity, and store origin in a completely within-subjects design in which all twenty-seven conditions were presented to each subject. Not surprisingly with such a transparent design, they found significant main effects of all three variables on quality ratings of sweaters with price accounting for the majority of explained variance. Contrary results were obtained by Jacoby, Olson and Haddock (1971) who manipulated the presence of information about price, brand name, and product composition on a between-subjects basis. No main effect of price on quality perceptions of beer was obtained from this less demand prone design. Interesting, however, was the fact that in the product-composition-absent cells, significant differences were found between the three product samples (the only factor which was manipulated on a within-subjects basis). In other words, where subjects were presented three samples of beer (which were identical) and were told that the samples were different brands, they acted as a demand bias explanation would assert and obligingly varied their ratings of three identical samples.

Nearly all repetition experiments have varied repetition on a within-subjects basis. Zajonc *et al.* (1971) observed that the positive effects of repetition on liking were dependent on the *relative* levels of exposure and not the *absolute* levels. The positive effects of repetition did not vary between separate experiments in which within-subject exposure levels were varied from 0, 1, 3, to 9; 0, 1, 3, 9, to 27; or 0, 3, 9, 27, to 81. However, when exposure was varied in a fourth experiment in which a subject viewed only one level of repetition, and could not compare repetition levels, no effect of repetition on liking was found. The fact that, in contrast to Zajonc's consistent positive effect of repetition in within-subject designs, the only experiment using a between-subjects design⁴ showed no effect of repetition provides more support for a demand characteristics explanation.

Although within-subjects designs will continue to

⁴ Zajonc *et al.* (1971) did find a positive effect of past word frequency (based on Thorndike-Lorge word counts) on liking of fruits and vegetables where frequency was a between-subjects treatment. However, the causal direction of this observed relationship is questionable since the question of whether more familiar things are more liked or more liked things are more familiar remains unresolved.

have an important role in consumer behavior research, their use appears to make demand bias a more plausible rival explanation. Where within-subjects designs seem advised, potential demand bias might be reduced by the use of Latin Square or incomplete block designs which present subjects with only some of the treatment combinations.

The second experimental design issue with respect to demand characteristics involves the use of non-experimental partitioning of subjects on the basis of some characteristic instead of experimentally manipulating that subject factor. This design usually groups subjects on some variable as a result of their performance or answers to some pre- or post-test. Although the non-experimental nature of such an independent variable drastically weakens the ability to infer causation, this design is often used in studies of personality in consumer behavior due to the difficulty in manipulating integral individual characteristics (see Kassarjian, 1971). Such non-experimentally manipulated individual variables may be correlated with demand-proneness. Variables such as intelligence, suspiciousness, independence, confidence, originality, idealism, impetuosity, empathy, and radicalism (Hampden-Turner, 1970) could act in such a manner. In instances where a trait cannot be satisfactorily manipulated experimentally, extra care must be taken to design the experimental measures and procedure to be as demand-free as possible.

Measurement of dependent variables

At least two aspects of measurement pose demand cue problems. First, measurement at the end of an experiment may be the cue that finally alerts the subject to the experimenter's hypothesis. This is especially true in attitude change research (Lana, 1959). Second, invalid measures may be affected more by demand cues than by the intended independent variables. The first problem can be handled by obscuring the connection between the experiment and the measurement. Techniques include using a confederate to ask the subject's reactions, pretending the measure is an unimportant afterthought, or positioning the post measurement as a pretest for a subsequent study.⁵ An alternative method to reduce measurement demand cues is the use of unobtrusive measures (Webb *et al.*, 1966). Although the latter method is difficult to implement, some success is possible (Ray and Sherrill, 1973).

Similarly, the measurement validity problem is also a difficult one. Although often less sensitive, dependent variables involving behavior are likely to be less demand prone. In some instances, physiological measures may be necessary to get valid measures of subjects

⁵ Rosenberg (1965) used this latter technique to effectively disguise measurement in an apparently successful attempt to decrease subject evaluation apprehension.

DEMAND ARTIFACTS IN CONSUMER RESEARCH

27

adopting artificial subject roles.⁶ Ideally, multiple measures can be used in order to both validly measure the intended construct (Heeler and Ray, 1972) and to accurately and unambiguously measure adopted roles.

Experimental procedures

Several procedures might be used to reduce demand characteristics. Most prominent is the use of deception. In many cases, the use of an appropriate cover study can help the subject accept unusual independent treatments and measurement conditions. In this regard, Aronson and Carlsmith (1968) advocate the experimenter's "tying up loose ends" of an experimental procedure himself rather than leaving such an ambiguous task to the subject. They similarly argue to risk experimenter bias by using a "live" experimenter to best present a convincing experimental cover to subjects.

Ironically, the use of deception to avoid demand bias can result in a heightening of such bias if the experimenter cover is insufficient. Thus, a genuine attempt must be made to measure the effectiveness of any disguise in the post-experimental inquiry. In a price-quality study where post-taste quality ratings of butter and margarine was the dependent variable, Valenzi and Andrews (1971) tried to disguise the true purpose by telling subjects that the question under study concerned whether mouth rinsing between tastes affected the taste sense. The price information was justified by stating that past subjects had asked for cost information. Despite what was probably the most elaborate disguise that has been reported in price-quality experiments, twenty-one of forty-six subjects subsequently guessed the true purpose of the investigation. In a conformity experiment, Venkatesan (1966) used the post-experimental inquiry and dropped from the analysis subjects who had correctly deciphered the experimental disguise and guessed the true hypothesis. Similar use of the post-inquiry to eliminate suspicious subjects has been reported in experiments about cognitive dissonance and post-purchase behavior (e.g. Mittelstaedt, 1969).

In an advertising repetition experiment, Sawyer (1973; Ray and Sawyer, 1971) told women shopper subjects that he was testing a computerized, in-home shopping system where products would be demonstrated and sold direct from store warehouses. Post-treatment measurement revealed that the subjects accepted the cover in an involved manner, and the disguise provided an acceptable excuse for both the presentation of repeated product advertisements and subsequent subject reactions to them. Heteromethod replications should attempt to vary the experimental disguise lest the cover, even though undeciphered, interact with some treatment variable. For example, the "Shopping of the Future" demonstration might have

accounted for an interaction between ad repetition and product type in Ray and Sawyer's (1971) study. The fact that repeated exposures were more effective in increasing ad recall and brand purchase intention for convenience goods than for shopping goods could have resulted if subjects perceived the futuristic shopping method as more appropriate for the former than the latter.

Given some of the problems associated with the use of deception, experimenters ought also to consider whether demand bias might be better alleviated with a procedure involving no deception. In attitude change research which *a priori* seems to carry the greatest justification for deception, Papageorgis (1968) found that whether subjects were deceived or forewarned about the purpose of a persuasion experiment was often not a crucial determinant of subsequent attitude change. He concluded that, in some instances, deception may cause more problems (including demand artifacts) than it solves. The assumption of greater external validity of experiments which attempt to conceal the fact that a message is designed to persuade the subjects has been questioned by McGuire (1969). He pointed out that natural persuasion attempts such as advertising are definitely not naively viewed as non-persuasive in intent. The most important aspect to disguise in a persuasion experiment is not the persuasive message intent but the fact that the subject's persuasiveness is of interest to the experimenter.

It has been suggested that straightforward role playing might replace the normal experimenter-subject relationship in experimentation and could especially obviate the need for deception (e.g., Kelman, 1967). Although role playing might be appropriate in some instances, most reviews of the evidence are unfavorable in terms of both internal and external validity (Miller, 1972). A major problem again involves the fact that a subject may be unwilling or unable to accurately predict his behavior. Especially confounding is the tendency for active role players to adopt the ambiguous apprehensive subject role. Also while simple effects may be replicated, more complex interactions are often not (e.g., Willis and Willis, 1969).

A procedure related to deception that is often suggested to reduce demand effects is the use of natural environments or complete field experimental settings. Consumer research issues provide many excellent opportunities to use natural settings. Although it is not likely to be a common alternative to the laboratory for several cost and control reasons, replication of results from lab procedures with hypothesized demand characteristic in less demand-prone field settings is impressive. For example, Zajonc's positive effects of repetition in the lab have been replicated in the field (e.g., Zajonc and Rajecki, 1969; Rajecki and Wolfson, 1972); however, neutral and negative effects have also been found (Strong, 1972). No further discussion is needed here

⁶ Note, however, that even Galvanic Skin Response is not immune to demand bias (Orne, 1969).

except to caution that natural or field experiments are not a panacea for the demand characteristics problem (e.g., Rosen, 1970; Roethlisberger and Dickson, 1939).

As mentioned earlier, a particularly difficult subject reaction to demand cues to measure and control is the apprehensive role. Any procedures that reduce subject apprehensiveness might be helpful. Low emphasis on experimenter status, non-experimenter control over rewards and punishments, obvious subject response anonymity, as well as disguised measurement can help alleviate subject apprehension. For example, it should be clear to the subject that no negative evaluations will be attached to those who actively try to maximize their leftover money in pricing experiments where subjects are allowed to keep the remainder of an original fixed sum in "buying" situations (e.g., Pessemier, 1959). However, care must be made to permit a too low level of subject involvement lest the whole experience appear too trivial and all experimental realism is lost.

A final procedural suggestion is to include the previously discussed post-experimental inquiry and non-experiment as standard quasi-control groups (Orne, 1969). Like other pretests on questionnaires and other measures, it is best to use either or both to pretest for demand characteristics before the time and effort of an entire experiment is invested. The use of these pretests will make unnecessary the later rejection of apparently confounded data.

CONCLUSION

It should now be obvious that a researcher must carefully examine an experimental laboratory procedure for the possibility of confounding demand artifacts. This is true whether the experiment in question is one being devised for the first time, being reviewed for interpretation of results after completion, or being considered for revision in replication. Such examination is also necessary whether experimental results confirm or disconfirm the experimenter's hypothesis and whether the experiment in question was conducted by the researcher himself or someone else. A general awareness and commitment is more important than any specific methodology. It is hoped that, in addition to specific suggestions, this paper has provided both an awareness of the potential importance of demand artifacts in laboratory experiments in consumer research and a motivation to try to detect and reduce such cues.

Finally, it is hoped that this paper will not be perceived as an attempt to disavow the use of laboratory experimentation in consumer research. Certainly, artifactual problems such as demand characteristics are not unique to laboratory experimentation (e.g., Wiseman, 1972). Instead, attention to demand characteristics should result in a greater confidence in lab methods and particular results either by resolving speculative criticism or by correcting for demand-bias results and

methods. Moreover, this paper should not cause consumer researchers to retreat to a womb of inactive theoretical research and sole concentration on demand characteristics and other methodological artifacts. Rather, it is hoped that this paper will lead to an integration of the suggested techniques of detection and control into on-going and new research.

McGuire (1969) described three stages in the "life cycle" of an experimental artifact: ignorance, coping, and exploitation. His discussion demonstrated how concern with demand characteristics, in addition to helping to resolve any doubts about the actual source of obtained effects, can lead to better understanding of the underlying theoretical processes. Thus, it should be realized that, like other methodological problems, proper concern with demand artifacts in laboratory experimentation is the essence of the scientific method and ethic. No greater argument for the increased attention of consumer researchers can be offered.

REFERENCES

Adair, J. G. and B. S. Schachter. "To Cooperate or to Look Good?: The Subjects' and Experimenters' Perceptions of Each Others' Intentions," *Journal of Experimental Social Psychology*, 8 (January 1972), 74-85.

Agyris, C. "Some Unintended Consequences of Rigorous Research," *Psychological Bulletin*, 70 (September 1968), 185-197.

Anderson, L. K., J. R. Taylor, and R. J. Holloway. "The Consumer and his Alternatives: An Experimental Approach," *Journal of Marketing Research*, 3 (February, 1966), 62-67.

Andrews, I. R. and E. R. Valenzi. "Combining Price, Brand Name, and Store Cues to Form An Impression of Product Quality," *Proceedings of the 79th Convention of the American Psychological Association*, 6 (1971), 649-650.

Aronson, E. and J. M. Carlsmith. "Experimentation in Social Psychology," in Gardner Lindzey and Eliot Aronson, *The Handbook of Social Psychology*, Vol. II. Reading: Addison Wesley, 1968, 1-79.

Becknell, J. C., W. R. Wilson, and J. C. Baird. "The Effect of Frequency of Presentation on the Choice of Non-sense Syllables," *Journal of Psychology*, 56 (July 1963), 165-170.

Bem, D. J. "Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena," *Psychological Review*, 74 (May 1967), 183-200.

Brehm, J. W. *A Theory of Psychological Reactance*. New York: Academic Press, 1966.

Brehm, J. W. and A. R. Cohen. "Re-Evaluation of Choice Alternatives as a Function of Their Number and Qualitative Similarity," *Journal of Abnormal and Social Psychology*, 58 (May 1959), 373-378.

Brock, T. C. and I. A. Becker. "Debriefing and Susceptibility to Subsequent Experimental Manipulation," *Journal of Experimental Social Psychology*, 20 (July 1966), 314-323.

Burgess, T. D. G. and J. M. Sales. "Attitudinal Effects of

DEMAND ARTIFACTS IN CONSUMER RESEARCH

'Mere Exposure': A Reevaluation," *Journal of Experimental Social Psychology*, 7 (July 1971), 461-472.

Campbell, D. T. and J. C. Stanley. "Experimental and Quasi-Experimental Designs for Research on Teaching," in N. L. Gage (Ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally, 1963.

Craig, C. S., J. F. Engel, and W. W. Talarzyk. "Consumer Decision-making: On the Importance of Price," in David M. Gardner, *Proceedings, Second Annual Conference of the Association for Consumer Research*, 1971, 243-255.

Deering, B. J. and J. Jacoby. "Price Intervals and Individual Price Limits as Determinants of Product Evaluation and Selection," in M. Venkatesan, *Proceedings, Third Annual Conference of the Association for Consumer Research*, 1972, 145-166.

Fillenbaum, S. "Prior Deception and Subsequent Experimental Performance: The 'Faithful' Subject," *Journal of Personality and Social Psychology*, 4 (November 1966), 533-537.

Golding, S. L. and E. Lichtenstein. "Confession of Awareness and Prior Knowledge of Deception as a Function of Interview Set and Approval Motivation," *Journal of Personality and Social Psychology*, 14 (1970), 213-223.

Grice, G. R. "Dependence of Empirical Laws Upon the Source of Experimental Variation," *Psychological Bulletin*, 66 (December 1966), 488-498.

Hampden-Turner, C. *Radical Man: The Process of Psycho-Social Development*. Cambridge: Schenkman, 1970.

Heeler, R. M. "A Laboratory Investigation of Inter-Related Effects of Mixed Media, Multiple Copy, and Multiple Insertions in Advertising Campaigns." Unpublished doctoral dissertation, Stanford University, 1972.

Heeler, R. M. and M. L. Ray. "Measure Validation in Marketing," *Journal of Marketing Research*, 99 (November 1972), 361-370.

Horowitz, I. A. and B. H. Rothschild. "Conformity as a Function of Deception and Role Playing," *Journal of Personality and Social Psychology*, 14 (March 1970), 224-226.

Jacoby, J., J. C. Olson, and R. A. Haddock. "Price, Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality," *Journal of Applied Psychology*, 55 (December, 1971), 570-579.

Kassarjian, H. H. "Personality and Consumer Behavior: A Review," *Journal of Marketing Research*, 8 (November, 1971), 409-418.

Kelman, H. C. "Human Uses of Human Subjects: The Problem of Deception in Social Psychological Experiments," *Psychological Bulletin*, 67 (January 1967), 1-11.

Lana, R. "Pretest Sensitization," in Robert Rosenthal and Ralph L. Rosnow, *Artifact in Behavioral Research*. New York: Academic Press, 1969, 119-141.

Leavitt, H. J. "A Note on Some Experimental Findings about the Meanings of Price," *Journal of Business*, 27 (July 1954), 205-210.

Levy, L. Y. "Awareness, Learning, and the Beneficent Subject as Expert Witness," *Journal of Personality and Social Psychology*, 6 (No. 3, 1967), 365-370.

LoSciuto, L. and R. Perloff. "Influence of Product Preference on Dissonance Reduction," *Journal of Marketing Research*, 4 (August, 1967), 286-290.

LoSciuto, L. A., L. H. Strassman, and W. D. Wells. "Advertising Weight and the Reward Value of the Brand," *Journal of Advertising Research*, 7 (June 1967), 34-38.

Masling, J. "Role-Related Behavior of the Subject and Psychologist and its Effects upon Psychological Data," *Nebraska Symposium on Motivation*, 14 (1966), 67-103.

McConnell, J. D. "The Price-Quality Relationship in an Experimental Setting," *Journal of Marketing Research*, 5 (August 1968), 300-303.

_____. "The Alphabet and Price as Independent Variables: A Note on the Price-Quality Question," *Journal of Business*, 43 (October 1970), 448-451.

McGuire, W. J. "Suspiciousness of Experimenter's Intent," in Robert Rosenthal and Ralph L. Rosnow, *Artifact in Behavioral Research*. New York: Academic Press, 1969, 13-57.

Miller, A. G. "Role-Playing: An Alternative to Deception? A Review of the Evidence," *American Psychologist*, 27 (July 1972), 623-636.

Miller, S. J., M. B. Mazis, and P. L. Wright. "The Influence of Brand Ambiguity on Brand Attitude Development," *Journal of Marketing Research*, 8 (November 1971), 445-449.

Mittelstaedt, R. "A Dissonance Approach to Repeat Purchasing Behavior," *Journal of Marketing Research*, 6 (November, 1969), 444-446.

Olson, J. C. "Cue Properties of Price," paper presented at the 81st Annual Convention of the American Psychological Association, Montreal, 1973.

Orne, M. T. "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implication," *American Psychologist*, 17 (November 1962), 776-783.

_____. "Demand Characteristics and the Concept of Quasi-Controls," in Robert Rosenthal and Ralph L. Rosnow, *Artifact in Behavioral Research*. New York: Academic Press, 1969, 143-179.

Papageorgis, D. "Warning and Persuasion," *Psychological Bulletin*, 70 (October, 1968), 271-282.

Peacock, P. and H. L. Davis. "The Alphabet as an Independent Variable: A Reply to J. Douglas McConnell," *Journal of Business*, 43 (April 1970), 205-209.

Pessemier, E. A. "A New Way to Determine Buying Decisions," *Journal of Marketing*, 24 (October 1959), 41-46.

Peterson, R. A. "The Price-Perceived Quality Relationship: Experimental Evidence," *Journal of Marketing Research*, 7 (November 1970), 525-528.

Rajecki, D. W. and C. Wolfson. "The Rating of Materials Found in the Mailbox: Effects of Frequency of Receipt," *Public Opinion Quarterly*, 110-114.

Ray, M. L. and A. G. Sawyer. "A Laboratory Technique for Estimating the Repetition Function for Advertising Media Models," *Journal of Marketing Research*, 8 (February 1971), 20-29.

Ray, M. L. and P. N. Sherrill. "Unobtrusive Marketing Research Techniques," in Steuart H. Britt, *Marketing Handbook*. Chicago: Dartnell, 1973, 317-330.

Riecken, H. W. "A Program for Research on Experiments

in Social Psychology," in N. F. Washburne, *Decisions, Values, and Groups*, Vol. 2, New York: Pergamon Press, 1962.

Roethlisberger, F. J. and W. J. Dickson. *Management and the Worker*. Cambridge: Harvard University Press, 1939.

Rosen, N. A. "Demand Characteristics in a Field Experiment," *Journal of Applied Psychology*, 54 (April, 1970), 165-171.

Rosenberg, M. J. "When Dissonance Fails: On Eliminating Evaluation Apprehension from Attitude Measurement," *Journal of Personality and Social Psychology*, 1 (No. 1, 1965), 28-42.

_____. "The Conditions and Consequences of Evaluation Apprehension," in Robert Rosenthal and Ralph L. Rosnow, *Artifact in Behavioral Research*. New York: Academic Press, 1969, 280-349.

Rosenthal, R. and R. L. Rosnow. *Artifact in Behavioral Research*, New York: Academic Press, 1969.

Rosnow, R. L. and L. S. Aiken. "Mediation of Artifacts in Behavioral Research," *Journal of Experimental Social Psychology*, 9 (May 1973), 181-201.

Rosnow, R. L., B. E. Goodstadt, J. M. Suls, and A. G. Gitter. "More on the Social Psychology of the Experiment: When Compliance Turns to Self-Defense," *Journal of Personality and Social Psychology*, 27 (September 1973), 337-343.

Saegert, S., W. Swap, and R. B. Zajonc. "Exposure, Context and Interpersonal Attraction," *Journal of Personality and Social Psychology*, 25 (February 1973), 234-242.

Sawyer, A. G. "The Effects of Repetition of Refutational and Supportive Advertising Appeals," *Journal of Marketing Research*, 10 (February 1973), 23-33.

Sawyer, A. G. "Detecting Demand Characteristics in Laboratory Experiments in Consumer Research: The Case of Repetition-Affect Research." Paper presented at the Annual Meeting of the Association for Consumer Research, November, 8, 1974(a).

_____. "The Effects of Repetition: Conclusions and Suggestions about Experimental Laboratory Research," in G. David Hughes and Michael L. Ray, *Consumer/Buyer Information Processing*. Chapel Hill: University of North Carolina Press, 1974, 190-214(b).

Sigall, H., E. Aronson, and T. Van Hoose. "The Cooperative Subject: Myth or Reality?," *Journal of Experimental Social Psychology*, 6 (January 1970), 1-10.

Silverman, I. "Role-Related Behavior of Subjects in Laboratory Studies of Attitude Change." *Journal of Personality and Social Psychology*, 8 (April, 1968), 343-348.

Stang, D. J. "The Effects of 'Mere Exposure' on Learning and Affect," *Journal of Personality and Social Psychology*, 30 (November 1974), 647-653.

_____. "Methodological Factors in Mere Exposure Research," *Psychological Bulletin*, in press.

Strong, E. C. "The Effects of Repetition in Advertising: A Field Experiment." Unpublished doctoral dissertation, Stanford University, 1972.

Suedfeld, P., Y. M. Epstein, E. Buchanan, and P. B. Landon. "Effects of Set on the 'Effects of Mere Exposure,'" *Journal of Personality and Social Psychology*, 17 (February 1971), 121-123.

Tybout, A. M. and G. Zaltman. "Ethics in Marketing Research: Their Practical Relevance," *Journal of Marketing Research* 11 (November 1974), 357-368.

Tull, D. S., R. A. Boring, and M. H. Gonsior. "A Note on the Relationship of Price and Imputed Quality," *Journal of Business*, 37 (April 1964), 186-191.

Valenzi, E. R. and I. R. Andrews. "Effect of Price Information on Product Quality Ratings," *Journal of Applied Psychology*, 55 (February 1971), 87-91.

Venkatesan, M. "Experimental Study of Consumer Behavior Conformity and Independence," *Journal of Marketing Research*, 3 (November 1966), 384-387.

Webb, E. J., D. T. Campbell, R. F. Schwartz, and L. Sechrist. *Unobtrusive Measures: Nonreactive Measures in the Social Sciences*. Chicago: Rand McNally, 1966.

Weber, S. J. and T. P. Cook. "Subject Effects in Laboratory Research: An Examination of Subject Roles, Demand Characteristics, and Valid Inference," *Psychological Bulletin*, 77 (April 1972), 273-295.

Willis, R. H. and Y. A. Willis, "Role Playing vs. Deception: An Experimental Comparison," *Journal of Personality and Social Psychology*, 16 (November 1970), 472-477.

Wiseman, F. "Methodological Bias in Public Opinion Surveys," *Public Opinion Quarterly*, 36 (Spring 1972), 105-108.

Zajonc, R. B. "Attitudinal Effects of Mere Exposure," *Journal of Personality and Social Psychology*, Monograph Supplement, 9 (June 1968), 1-27.

Zajonc, R. B. and D. W. Rajecki. "Exposure and Affect: A Field Experiment," *Psychonomic Science*, 17 (November 1969), 216-217.

Zajonc, R. B., W. C. Swap, A. A. Harrison, and P. Roberts. "Limiting Conditions of the Exposure Effect: Satiation and Relativity," *Journal of Personality and Social Psychology*, 18 (June 1971), 384-391.

THE JOURNAL OF CONSUMER RESEARCH

Copyright of Journal of Consumer Research is the property of Journal of Consumer Research, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.